Bárbara Bordalejo

The Texts We See and the Works We Imagine: The Shift of Focus of Textual Scholarship in the Digital Age

Introduction

The process of editing a text is, in the first instance, an act of imagination. An editor who has collected materials, gathered evidence, and compared variants eventually has to decide what does it all mean, who will care about it and how to present it; but most importantly how those materials relate to each other. The answers to these questions are not in the documents that preserve versions of the texts, but in the minds of the scholars who have carefully studied the physical documents, their texts and the variant states of the text they represent. In this essay, I present my working definitions of the "text of the document", the "variant states of the text" and the "work", show how they relate to each other and how they have been affected by digital technologies or how they have arisen from them. I also conclude that while some concepts might remain unchanged from the days of print, others are fundamental only to born-digital texts.

At the beginning of the 80s, Jerome McGann and D.F. McKenzie, almost simultaneously but independently, questioned the textual critical establishment and its intentionalist view of editing.[2] Up to that point, first Fredson Bowers and later G. Thomas Tanselle, both following W.W. Greg, advocated copy-text editing and the recovery of authorial intention. For McGann and McKenzie, the text was a result of a series of physical processes that had to be taken into account, as well as explained, as part of the editorial process.

With the advent of digital editions, editors found that they could include an amount of information that would not have been possible in print where the only limitations were time and money. The first electronic editions, Kevin Kiernan's *Beowulf*; Peter Robinson's *The Wife of Bath's Prologue on CD-ROM*; Murray McGillivray's *Book of the Duch-*

[1] I would like to thank Paul Eggert for his comments on the first draft of this paper.

[2] J.J. McGann, *A Critique of Modern Textual Criticism*, Chicago, Chicago University Press, 1983; D.F. McKenzie, *Bibliography and the Sociology of Texts* (The Panizzi Lectures, 1985), Cambridge, Cambridge University Press, 1999.

ess; Viscomi, Essick and Eaves' *The Blake Archive*; McGann's *The Rossetti Archive;* Price's *The Whitman Archive*, rely heavily on images and present a series of transcriptions of documents. Editors, perhaps enamored of the new medium, stopped editing and began compiling vast amounts of images of documents with their respective transcriptions. This did not mean that the theory of textual editing had been set aside, but despite various editors considering the subject, the production of critical digital editions was left for some other occasion.[3]

Documents and texts

In leaving the production of critical editions for the future, editors have focused on the transcription of primary textual sources and the production of digital surrogates for those documents containing text. This has led Hans Walter Gabler to describe texts as the functions of documents:

> Digital editions must however, in their turn, and precisely in their 'otherness', derive bearings from their texts' native transmissions in the material medium. This is where recognizing the primacy of the document – meaning that texts are, logically, always functions of the documents transmitting them – becomes essential. It is exactly where, and when, the text is and remains separate from the material support of its transmission that the material parameters of that support need to be adjudicated as potential determinants for the digital edition. To see the text fundamentally as a function of the document helps to recognize afresh that in all transmission and all editing, texts are and, if properly recognized, always have been constructs from documents.[4]

Here Gabler characterizes the process of «constructing» the text, which he identifies with the editorial process of creating a critical edition, and he declares the primacy of documents before offering the idea that «texts are functions of documents». This is a very imprecise definition: documents are the material support for texts and do not cause texts to do anything. Only a reader can extract meaning from texts; each reader, depending

[3] Critical editions have continued to appear in print. Some of these have benefited from using digital methods for their realization, for example, the Nestle-Aland edition of the Greek New Testament or the *Editio Critica Maior*. See http://www.uni-muenster.de/INTF/Publications.html#Editio_Critica_Maior, accessed 13 September 2013. See also F. Rico, «Texto y textos en tiempos de crisis», *Medioevo romanzo*, 35 (2011), pp. 58-65: 64-65.

[4] H.W. Gabler, «Theorizing the Digital Scholarly Edition», *Literature Compass*, 7.2 (2010), pp. 43-56: 51.

on his or her background, might extract more or less meaning or might comprehend one or various levels of that meaning. Of course, one could, as Peter Robinson has done, recognize the merits of paying close attention to documents, while highlighting the dangers of the «flood of facsimile editions in digital form» that have already resulted.

Elena Pierazzo also gives great weight to documents and documentary editions.[5] Her exposure to Jane Austen's draft manuscripts have given shape to her understanding of the importance of diplomatic transcriptions. However, she takes this understanding to a new level in advocating that:

> [a] diplomatic edition is a non-objective, interpretative operation, then it follows that [it] is also a scholarly activity and can be justified on these grounds, in the same way that a critical edition can, with both presenting the scholarly and computational analysis of the chosen textual phenomena.[6]

From the above quotation, it becomes clear that Pierazzo has a particular agenda: having worked as part of the Jane Austen's Fiction Manuscripts, she finds herself in the position of having to prove that documentary editions are, at least, equal to critical editions.[7] In the process of doing this, she defines a

> ...new type of editorial object, the *documentary digital edition*, as the recording of as many features of the original document as are considered meaningful by the editors, displayed in all the ways the editors consider useful for the readers, including all the tools necessary to achieve such a purpose.[8]

It is unclear whether Pierazzo thinks that the innovation came before or after the Jane Austen's Fiction Manuscripts edition, but given the absence of any other examples in her article it appears she believes this is the first edition in this category.

[5] E. Pierazzo, «A Rationale of Digital Documentary Editions», *Literary and Linguistic Computing,* 26 (2011) pp. 463-477.

[6] Ibidem, p. 472.

[7] Jane Austen's Fiction Manuscripts: http://www.janeausten.ac.uk/index.html, accessed 16 September 2013. The assertion of the value of documentary editions is beating a dead horse: editions have different aims and serve different purposes, they all present (and always have presented) a degree of subjectivity. Certain types of edition are not inherently superior to others, so there is no need to try to «vindicate» any type. Textual scholars know this.

[8] Pierazzo, «A Rationale of Digital Documentary Editions», p. 475.

The text of the document

A document that has been inscribed with text has no function unless a reading agent is present.[9] Once this reading agent is present, meaning can be extracted from the text preserved in the document. For most competent readers, the action of reading is so habitual and occurs so quickly that they fail to notice all the underlying processes that their experiences bring into the equation.

What I call the text of the document is the totality of the text as preserved on its physical support. Scholars with experience with the transcription of primary sources will immediately understand that I am referring to all the meaningful marks on the page made by someone with the intention of communicating something. The meaningful, intentional marks that are not script, that is that do not contribute to the representation of words or pauses (punctuation), are included as part of the text of the document.[10] To make it really clear, any indications as to which text might be considered erased or what needs to be included, marks that suggest a change in order or any other meaningful signs on the page, are part of this text of the document. Thus, the phrase the «text of the document» refers to the complete sequence of marks present in the document, independently of whether these represent a complete, meaningful text in itself. That is: the reader sees a sequence of letters, occurring in various places in relation to each other (perhaps between the lines or within the margins) and carrying various markings (perhaps underdottings or strikethroughs).

The text(s) in a document: the variant states of the text

The text of the document, as described above, can be interpreted to mean more than just one variant text. In this way, a single document contains two or more variant states. I have stated elsewhere that:

[9] This agent could be human or machine. For the purposes of the latter the text should have to be encoded to be read. For the purposes of this essay, when I refer to a reading agent, I am talking about a person who both has a high-level understanding of the language of the text and who has reached an adult level of reading comprehension. Such a person, I call a competent reader.

[10] A document can present marks that are the result of accidents (ink splatters, stains) or exposure (dust). These unintentional marks are not part of the text of the document.

The reader understands the marks present in the text of the document as meaningful and constructs one or more specific senses from them. Where more than one sense can be constructed from the text of the document, I refer to these as the «variant states of the text», or as the constructed texts. I deliberately avoid the use of the phrase the «text of the work», as this is a completely different concept that refers specifically to an abstract concept of «the work».[11]

So in the process of interpreting the marks on the document, often without even realizing it, a reader can abstract several variant states of the text from a single document. This act of abstraction, as I explained before, becomes much clearer when one is in the process of transcribing a text. This is due to the difference in nature, as well as in speed, that transcription entails. To synthesize, the text of the document is made up of *the meaningful marks on the page*. The variant states of the text *are extrapolated by an individual* (a reader) who, consciously or unconsciously, makes decisions about how to interpret the text of the document.

What is a work?

While writing on the *Divine Comedy*'s encoding system, I stated that I never used the phrase the «text of the work» because I might only use that concept in reference to the abstract (some might call ideal) concept of the work, a concept linked to ideas expressed, on more than one occasion, by Tanselle.

I owe a great debt to Tanselle's writings and his teaching. However, there are some areas in which I cannot completely agree with him. These points of connection and contrast relate particularly to Tanselle's concept of work, as presented in *A Rationale of Textual Criticism*:

Even the most unsophisticated readers have sometimes decided that a particular formation of letters or sequence of words – apparently meaningless in the language being used or inappropriate in context – is a «typographical error» or a «slip of the pen», and in so doing they have perhaps faced more aesthetic issues than they knew. They were first of all showing that they wished to understand what was intended by someone else... Then they were implicitly claiming that they had been able to locate the real work – the real statement, though not neces-

[11] B. Bordalejo, «The Encoding System», in Prue Shaw, ed., *Dante Alighieri. Commedia. A Digital Edition*, Birmingham and Florence, Scholarly Digital Editions and SISMEL-Edizioni del Galuzzo, 2010.

sarily the real or only meaning – hovering somehow behind the physical text, which had served as an occasionally unreliable, but always indispensable, guide to it. They were also recognizing that what they had recovered (or attempted to recover) was not simply someone's thought but the actual expression of those thoughts – that (whatever we take the relation between thought and language to be) verbal works or statements are thoughts employing particular arrangements of words as their ultimate medium.[12]

This concept is now recognized, whether other scholars are in agreement with it or not, as a classic one in Anglo-American textual criticism, a field in which considering the «work» always leads to the discussion of Tanselle's ideas. These discussions centre on the question of the materiality of the work which the above quotation states is not really material (although material traces are used in its construction), but rather the expression of someone's thoughts. Further, in imagining a correction, one is attempting to recover an expression which only existed before those thoughts were ever put onto paper. It is no wonder that Tanselle has been referred to as an idealist. Despite the clear materiality of the documents and his insistence in highlighting them as the only traces left of the work, the work itself appears not to have any materiality. In this sense, for Tanselle, the work is an abstract concept, an idea.

I tend to agree with Tanselle's definition as I understand it, but I do so with reservations. Tanselle is thinking about what could only be described as authorial works, probably literary and most likely post-Romantic. For most post-Romantic literature, one requires only a relatively small leap, a small act of imagination, to link texts and works, and to recover authorial intention from a multitude of documents. This abundance of drafts, printer's copies, separate magazine publications, proofs, personal diaries, letters, notes, etc, present an ideal environment for the possible restoration of authorial intention (the editorial position generally advocated by Tanselle). In comparison, pre-Romantic texts usually present a much less varied and rich set of documents. It is my experience with pre-Romantic works which makes me feel uncomfortable with Tanselle's definition, not because it invalidates his concept of the work (it doesn't), but because it implies that the concept must be put to use in the recovery of authorial intention; a recovery that is not always feasible with older texts.[13]

[12] G.T. Tanselle, *A Rationale of Textual Criticism*, Philadelphia, University of Pennsylvania Press, 1989, pp. 14-15.

[13] Although final authorial intention could be applied to any text of any time, it might be more easily carried out with 19th and early 20th centuries text.

Tanselle takes these notions, of work and document even farther to create a clear separation between the concepts of the «text of the work» and the «texts of documents». I have never have had a particular problem with the phrase the «text of the work» when it is used to describe the perfected sequence of words that were meant to be, let us say, *David Copperfield*. I have a problem with the idea of the recoverability of the texts of the work in all circumstances and with the idea that this is an adequate approach to edit pre-Romantic texts. But this is the subject of a different discussion. The most important conclusion that is drawn from the dichotomy is that the text of any document is different from the text of the work, and so, by my own definition, each of the variant states of the text is also different from the text of the work.

A substantial amount of my textual critical research relates to works which are much older than those ordinarily considered by Tanselle. Most of those works are authorial: Dante's *Divine Comedy*, Chaucer's *Canterbury Tales*, Boccaccio's *Teseida*. Some, however, like 15th century Castilian *Cancioneros* do not easily fall into that category.[14] It is not that Tanselle's notion of work ceases being effective, but rather that the evidence preserved in the extant documents lets us down and would never allow us to pursue the evasive intentionality which seems to be the final objective of the Tansellian system. In this, I agree with Paul Eggert when he states that:

> Shillingsburg deviates from McLaverty here but follows him (and also Tanselle) on another level when he argues that a work is only implied by physical manifestations of it; it is not identical with any of them. The reader infers the existence of the work and its text (or the versions of it and their texts), perhaps making allowances for any errors believed to be present. This text of the work is a 'Conceptual Text' since it is not materially witnessed: hence the traditional need for editorial action to recover it.[15]

Although, for some texts, the recovery of authorial intention might be possible, desirable or even necessary, the concept of work cannot just be constructed as part of an intentionalist agenda. If it were conceived only in this way, then its use would be very limited.

[14] To a lesser extent, I have explored other textual histories or studied the transmission of texts that not only do not have the benefit of having a single author, but that have developed during the course of many years: Middle Egyptian texts or the text of the Greek New Testament.

[15] P. Eggert, *Securing the Past*, Cambridge, Cambridge University Press, 2009, p. 229.

My background and my understanding of how texts function influence the way in which I understand the work. They force me to phrase the concept of work slightly differently: *the work is a conception in the mind of an author at a particular point in time that serves as a minimal denominator to identify its remaining physical manifestations*. If we take it in parts, «a concept in the mind of the author» is more akin to a general idea than to a series of sentences that form the totality of a novel, a poem or any other kind of text. When I refer to «a particular point in time» I bring in a time dimension (also considered by Eggert), which is very important in understanding that the creative process is not necessarily inspirational and instantaneous, but that it can be the result of years of crafting, molding and reshaping the structure of language. The idea of the «minimal denominator» is essential to my concept, because it is what marks the fact that the function of the work is classificatory in that it allows us to recognize different instances of one thing.

It might be possible to argue that, at some point, the author thought of a sentence and that this was later transcribed onto paper, but such a sentence might have been modified later by the same author (as it can often be seen in authorial drafts). The creative process in literature is not likely to be that an author suddenly conceived an idea that starts with the words «Of man's first disobedience...» and continued in absolute order to «Through Eden took their solitary way». It would be ludicrous to think that an author is just a «transcriber» of his or her own text. Perhaps some of Coleridge's contemporaries believed his account of the creation of «Kubla Khan»; few readers now do.[16]

In my concept, as much as in Tanselle's, the work, or even better, fragments of the work, are something that cannot be found in the realm of the material, not because they are not material, but because their materiality is so sudden and so fleeting that we no longer have them. Traces of the work, evidence of its existence, can be found in the documents and the texts they hold, but the work itself is none of these instances while, at the same time, is somehow present in all of them. To some observers, this might appear as «idealism». But in my definition, the function of the concept of work is to permit the recognition of its manifestations in the physical world, the recognition of the texts that are material expres-

[16] Even Edgar Allan Poe took to mocking this notion of creation in his essay «The Philosophy of Composition», first published in 1846. The essay is a fictional account of the process of composition of *The Raven*, which had been published in 1845.

sions of the work. The concept of work, defined as series of phenomena which occurred during fleeting moments in time, helps to build the bridge for that act of imagination that is editing.

The concept of work outlined above has served me very well as a textual critic and an editor. I use it particularly to shape my understanding of textual traditions, to make decisions as to which texts are to be compared with which others and to refer to the works that I am studying. Because in the world of textual transmission, our acknowledgement of a text being witness to a particular work is what defines whether that text belongs or not to a textual tradition.

The limits of the concept of work

When discussing my own concept of the work as a minimum denominator to establish a text as part of a textual tradition, I stated that I do not agree that one must always attempt in all circumstances to recover the work as the author's concept. For many works, the recovery of authorial intention is so unlikely as to make the task simply unworthy of any effort. In those cases in which the documentary evidence is so scarce, so far removed from the author or authors; one is better off employing the concept of the work as minimum denominator to relate existing documents to one another in an attempt to understand how they relate to each other and how they evolved.

There are many factors that affect an edition, and the concept of work is just one of them. For example, for a critical edition of the *Canterbury Tales*, I would follow new-stemmatic principles, and build a text that is the latest witness of the textual tradition,[17] one that calls attention to places of variation and explains both the relationships between different witnesses and the evolution of the text. In my work for the CantApp, an application for mobile devices, I have a very different objective in mind. The CantApp requires a reading text aimed at the general public,

[17] Peter Shillingsburg in his article «Literary Documents, Texts and Works Represented Digitally», in this collection, makes a similar point when he states that: «First, remember that print scholarly editions always produced new texts: sometimes accurate lexical reiterations of historical texts and sometimes eclectically edited new texts. In either case, the reading text was a new text, not an old one; and it was bolstered or surrounded by historical introductions and an apparatus of textual materials from alternative texts, which among other things attempted to indicate what was new and what was old about the newly produced scholarly edited text.»

but specially to undergraduate students. For the CantApp, I am creating a reading text that enhances the literary experience of the *Tales*. While producing editions such as these, I still hold the notion of the work that I have highlighted above and which has helped me to decide, for example, about canonicity and exclusion in reference to the tales and links that should be part of an edition of Chaucer's *Canterbury Tales*.

What is an editor?

The document-centered perspective, as championed by Gabler and Pierazzo, in digital editing is so pervasive that it has led some editors to exalt the document as if editing not only begins with the document, but ends with it too.

For Gabler the editor becomes a «facilitator»: that is, a mediator and a guide to the document.[18] Gabler is specifically referring to complex draft materials (as he does in other sections of his article), in which case the use of the word facilitator might appear less dissonant. However, for documents which are easily read by any competent reader, one hardly needs a facilitator. To talk about the editor as a facilitator is to diminish the importance of the work editors do. It is true that editors might make it possible for scholars and non-academics alike to read texts and to understand the complexities of textual transmission, but when an editor presents an edition after working on a text for an extended period of time and analyzing it from different perspectives he or she is really presenting a hypothesis that might be the result of years of research, analysis and reflection. The term facilitator falls short of describing the work that goes into the production even of the simplest type of edition. Although to study texts, particularly with the aid of computers, we might have to transcribe them, the final objective of textual critics is to understand the series of historical phenomena linked to the different variant states of the text, to offer precise accounts of these, and to formulate rational hypotheses to explain them.

[18] H.W. Gabler, «Theorizing the Digital Scholarly Edition», p. 52. The term «facilitator» is (sadly) repeated by Siemens *et al*. when discussing the so-called social edition, «Toward Modeling the *Social* Edition», *Literary and Linguistic Computing*, 27 (2012), p. 453.

Editing texts in the digital age

Many commentators on digital editions have asserted that scholarly editions in the digital medium are fundamentally different from those of the print age. Pierazzo states:

> It is the argument of this article that editions as we know them from print culture are substantially different from the ones we find in a digital medium.[19]

I argue that indeed we will see fundamentally different editions in the digital age. However, they will not be different from print editions for the reasons given by Pierazzo. According to her (and also Kiernan and Gabler), digital editions are different in their exclusive focus on the documents, while leaving aside the production of edited texts.[20] The digital medium has enabled this shift of focus, by breaking down the space restrictions usually linked to the production of printed editions. However we have had facsimile editions, often with transcriptions (sometimes, very elaborate) beside images, since the mid-nineteenth century. Nothing new here. The other areas which Pierazzo identifies as unique to the digital medium are the use of complex computer encoding, particularly using the TEI Guidelines, and the ability to record multiple phenomena (names, locations within the page, intricate transcriptions) within the one transcription. All this before concluding:

> Printed and digital editions may have the same function, namely to make a given text available to an audience, but the way they have to be prepared, the kind of questions the editor needs to answer, and ultimately their very natures are substantially, if not ontologically, different.[21]

Here Pierazzo confuses digital humanities questions with textual critical issues. Processing texts using digital methods requires a different preparation of the material because they need to be coded to be used with computers. This is true whether the final result is a digital edition or a printed edition. The questions that Pierazzo poses and that an editor

[19] Pierazzo, «A Rationale of Digital Documentary Editions», p. 463.

[20] K. Kiernan, «Digital facsimiles in Editing», in L. Burnard, K. O'Brien O'Keeffe and J. Unsworth, eds., *Electronic Textual Editing*, New York, Modern Language Association of America, 2006.

[21] Pierazzo, «A Rationale of Digital Documentary Editions», p. 475.

needs to answer have to do with the use of computers as tools to produce the edition. Again, this is true whether the final edition gets printed or published digitally. Textual scholars have always recorded fine detail, within transcriptions, in introductions, notes, tables and indices. One may argue that the TEI system allows these to be done more efficiently; but an increase of efficiency is not a revolution.

Daniel O'Donnell, in private conversation, has stated how puzzling he finds the lack of discussion of Digital Humanities subjects at the conference of the Society for Textual Scholarship. This should not be surprising. When Tanselle states that printed and digital editions «are not ontologically different» and that «their conceptual status of the text in each case is identical»,[22] he is right: the subject has not been changed by the methods we are using for our research. Even among textual scholars who are very involved in the production of digital editions, the subjects of discussion within the field are very much the same as they have always been: either theoretical (as this article and the ones that accompany in this special section) or practical (relating to the difficulties presented by particular texts, documents and authors). Textual scholars continue to use more or less the same methodologies and approaches that they had in the past. The main difference is how much easier some of those tasks have become. If we think, for example, about collation and the classification of variants, both activities can be carried out faster and with greater efficiency thanks to the computer but, as activities, they are not fundamentally different from those carried out before the digital era. This is not to say that nothing is different and although from a theoretical perspective we continue to face the same issues, there is one aspect of our work that is changing due to the use of computers: the status of copyright. For digital editions, and indeed all digital work, to continue and thrive, we need to encourage new forms of licensing that leave behind the nineteenth century notions of copyright and authorship. All editorial work (but also non-editorial work), should be licensed under the creative commons Attribution 3.0 unported.[23] This is the single difference, and a fundamental difference, between digital editions and its printed counterparts. The former

[22] G.T. Tanselle, «Foreword» in Burnard, O'Brien O'Keeffe, Unsworth, eds., *Electronic Textual Editing*, p. 6.

[23] Or a version that supersedes this one in the future. Notice that this license does not have a commercial restriction, a very important feature in a world where both site advertisement (a single commercial add in an otherwise free site could bring about legal action if there is a non-commercial restriction in place) and lawyers abound.

require the freedom that creative commons grants: freedom to reuse, to modify, to upcycle.[24]

Conclusions

There is no doubt that texts can change over time and that they indeed change. Our notion of what relates them to a historical event or series of events is informed by the concept of the work as presented in this article. It is possible to create a wider ranging concept, one that would encompass the work and its manifestations in physical form, and would be closer to Eggert's definition of the work as a «regulative idea», but which cannot be completely equated with my own concept of work.

This paper started with the statement that the process of editing a text is, in the first instance, an act of imagination. It is an act of imagination because we need to bridge the gap between the words that we see physically marked on the surface of the material document and our idea of the work as conceived by one or many authors who wanted to convey a set of ideas at a particular point in time. The act of imagination occurs when we leap from the text we see to the work we imagine, when, after researching and carefully considering our options, we choose one among a multitude of variant readings to be included as part of the text we are about to present. G. Thomas Tanselle would say that we use our critical judgment to achieve this decision and truly some editors do exactly this. But the act that takes us from one to the other, from the variant readings to the edited text, is as much the result of careful thought as it is the result of an instinct trained by years of study: it is a leap of imagination.

Peter Shillingsburg
Literary Documents, Texts, and Works Represented Digitally

The advent of digital technologies for preservation and dissemination of texts has complicated rather than revolutionized textual scholarship on literary texts. Disputes about methods and goals for scholarly textual studies in the fields of bibliography, textual criticism, and scholarly

[24] «Upcycling» refers to the action of repurposing something, that might have been worthless before, to create a new item with a new and unforeseen value.